Multiple sclerosis retrovirus-like envelope gene: Role of the chromosome 20 insertion

Jezabel Varadé a,1, Marta García-Montojo b,1, Belén de la Hera a,1, Iris Camacho a, Mª. Ángel García-Martínez a, Rafael Arroyo b, Roberto Álvarez-Lafuente b,2, Elena Urcelay a,⁎

a Immunology Dept, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), 28040 Madrid, Spain
b Multiple Sclerosis Unit, Neurology Dept, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), 28040 Madrid, Spain

⁎ Corresponding author at: Immunology Department, Hospital Clínico San Carlos, IdISSC, Martin Lagos s/n, 28040 Madrid, Spain. Tel.: +34 91 330 3000x7171; fax: +34 91 330 3879.
1 Joint First Authors.
2 Equal contributors.

A R T I C L E   I N F O
Article history:
Received 9 December 2014
Received in revised form 9 February 2015
Accepted 12 February 2015
Available online 19 February 2015

Keywords:
Human endogenous retrovirus HERV-W
MSRV
Multiple sclerosis
Genetic susceptibility

A B S T R A C T

Background: The genetic basis involved in multiple sclerosis (MS) susceptibility was not completely revealed by genome-wide association studies. Part of it could lie in repetitive sequences, as those corresponding to human Endogenous Retroviruses (HERVs). Retrovirus-like particles were isolated from MS patients and the genome of the MS-associated retrovirus (MSRV) was the founder of the HERV-W family. We aimed to ascertain which chromosomal origin encodes the pathogenic ENV protein by genomic analysis of the HERV-W insertions.

Methods/results: In silico analyses allowed to uncover putative open reading frames containing the specific sequence previously reported for MSRV-like envelope (env) detection. Out of the 261 genomic insertions of HERV-W env only 9 copies harbor the specific primers and probe featuring MSRV-like env. The copy from chromosome 20 was further studied considering its size, a truncated homologue of the functional HERV-W env sequence encoding syncytin. High Resolution Melting analysis of this sequence identified two single nucleotide polymorphisms, subsequently genotyped by Taqman chemistry in 668 MS patients and 678 healthy controls.

No significant association of these polymorphisms with MS risk was evidenced. Transcriptional activity of this MSRV-like env copy was detected in peripheral blood mononuclear cells from patients and controls. RNA expression levels of chromosome 20-specific MSRV-like env did not show significant differences between MS patients and controls, neither were related to genotypes of the two mentioned polymorphisms.

Conclusions: The lack of association with MS risk of the identified polymorphisms together with the transcription results discard chromosome 20 as genomic origin of MSRV-like env.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Background

Multiple sclerosis (MS) is a chronic inflammatory and neurodegenerative disease of the central nervous system affecting young adults [1]. Its precise etiology is not currently ascertained, but the prevailing hypothesis accepts the trigger of poorly-defined environmental factors acting in genetically susceptible individuals. The genetic background of these patients has been thoroughly studied and genome-wide association studies (GWAS) have increased significantly the number of loci identified in MS predisposition [2,3]. Nonetheless, the overall genetic load revealed by this robust approach does not seem to fully explain previous estimations based on epidemiologic studies [4]. Neither the analysis of rare coding-region variants in known risk genes for autoimmune diseases clarified where the missing heritability lies [5]. In parallel, research on environmental factors has often pointed to different virus putatively related with this condition [6].

Retrovirus-like particles were observed in cell cultures of MS patients [7] and the genome of the MS-associated retrovirus, MSRV, was the founder member of a human Endogenous Retrovirus family, HERV-W [8]. Some time ago the repetitive sequences within the human genome were considered “junk DNA”, but this issue has been widely debated [9,10] and recently, the ENCODE project has also questioned this gene-centric view of the genome [11]. Only about 1–3% of the genome corresponds to protein-coding sequences, while over 40% of the human DNA sequence holds repeated elements, and the endogenous retroviruses qualify for 8% of the total [12]. Most HERVs are defective for replication by the acquisition of stop codons, frameshift mutations and deletions [13]. Even so, MSRV virions have been reported in plasma and cerebrospinal fluid from MS patients and were associated with MS pathogenesis [8,14–16]. Moreover, potent immunopathogenic properties mediated by T cells have been described in vivo for MSRV retroviral particles from MS.
cultures [17] and the MSRV envelope (ENV) protein was claimed to be responsible for the observed damage [18,19]. Additionally, a recent study demonstrated that the protein MSRV ENV can be used instead of a mycobacterial lysate to induce Experimental Autoimmune Encephalomyelitis (EAE) [20].

The HERV-W retrovirus family is inserted into hundreds of genomic positions [21]; approximately half of them have been reduced to solitary LTRs through evolution and most of the remaining copies have rendered defective. The complete MSRV genome is around 7700-nucleotide long, with RUS, gag, pol, env, and U3R regions [22]. One functional HERV-W env copy from chromosome 7 codes the physiological syncytin [23]. Mameli and collaborators discovered a 12-nucleotide insertion in the transmembrane moiety of the MSRV env gene, which was absent from the syncytin genetic copy and might enable its discrimination [24]. Based on this insertion, the authors confirmed the link between MS and MSRV and concluded that syncytin expression does not differ in PBMCs from MS patients and controls [24]. Since the integration site of the true MSRV is still unknown, the unique sequence that can be defined as MSRV env is the one detected in extracellular virus particles. Therefore, all the HERV-W env DNA sequences with the features of the virionic MSRV env [i.e. the 12 bp insertion] must be defined as “MSRV-like” envs.

A yet unresolved issue is the genomic origin of the MSRV ENV protein. The insertion in chromosome X which contains a shortened version of the env gene has been regarded as a good candidate, given that this env sequence would encode a truncated syncytin-like protein of 475 amino acids [25]. In fact, a polymorphism described in that chromosome X copy was found associated with MS predisposition and with differential expression of MSRV [26]. In the present work, we aimed to further explore the chromosomal origin of the HERV-W ENV involved in MS pathogenesis, as the origin found in chromosome X is not necessarily the only one. GWAS do not analyze repetitive sequences from the genome in their search for susceptibility factors, therefore we hypothesized that part of the yet-undiscovered heritability could reside in these repetitive sequences.

2. Methods

2.1. Patients and controls

The Spanish case–control study included a total of 668 MS patients (64% females) and 678 healthy controls (57% females), with mean ages of 40 ± 9y and 41 ± 17y, respectively. All Caucasian participants were recruited from a single center, Hospital Clínico San Carlos, Madrid. MS diagnosis was established according to McDonald’s criteria [27]. MS patients [age at onset (mean ± SD): 26 ± 9y] were classified in relapsing remitting (81%), primary progressive (9%) and secondary progressive (10%). None of the control subjects reported first or second degree relatives with any immune-mediated disease. All subjects were recruited after written informed consent and the Ethics Committee from Hospital Clínico San Carlos approved this study. All research was conducted according to the principles expressed in the Declaration of Helsinki.

2.2. In silico analysis

The MSRV-env sequence (AF331500) was aligned with the human genome assembly GRCh37.p5 using BLAST (http://www.ensembl.org). The query results were analyzed by pDRAW32 (developed by AcaClone, http://www.acaclone.com) to identify putative open reading frames (ORFs) containing the specific primers and probe reported for MSRV env RNA detection [24]. Then, sequences were compared using a multiple sequence alignment tool (www.ebi.ac.uk/Tools/msa/clustalw2/) in order to find differential sequences for each insertion.

2.3. Expression analysis

RNA was extracted from peripheral blood mononuclear cells (PBMCs) from fresh blood by centrifugation in CPT tubes (Becton Dickinson), using the Qiamp RNA Blood Mini kit (Qiagen). After DNAse treatment, RNA sample volumes adjusted to 10 ng/μl were retrotranscribed using First Strand cDNA Synthesis kit (Roche Diagnostics, S.L. Barcelona, Spain) with OligoDT primers. Relative expression of MSRV env was measured by the 2e−ΔΔCt method. In short, quantitative real time PCR of the cDNA used a probe (GCCCTCTAACT GCCTCTCTGCT) and primers (forward: TATTTGGGAGGTTGGTGT; reverse: GCTAACAATCATTCTCAATAGG) designed using PRIMER3 v4.0 software (http://frodo.wi.mit.edu/) for MSRV-like env derived from chromosome 20, and Glucuronidase-beta as housekeeping gene (HKG, GUS-B assay from Applied Biosystems). Each amplification included samples with their no-RT controls to detect possible genomic DNA contamination. Assays for the detection of GUS-B and chromosome 20 HERV-W env were considered acceptable when: 1) the Ct for the HKG was lower than the Mean + 2 × S.D. of all samples; 2) no amplification was detected in the no-RT control; and 3) duplicates of each sample presented less than 2% of variability on Ct basis.

2.4. High Resolution Melting (HRM) analysis of MSRV-like env locus in chromosome 20 and genotyping

A PCR preparatory step was conducted for this chromosomal region (for primers and PCR amplification conditions see Supplementary Table 1). The PCR product containing the MSRV-env region of chromosome 20 was diluted 1/80 on sterile water, and then used as template to perform HRM analysis with a set of 12 pairs of primers that were designed using PRIMER3 v4.0 software [28]. Melting curves were discriminated in a 7900HT Fast Real-Time PCR System (Applied Biosystems, Foster City, CA, USA), according to the manufacturer's recommendations.
When a different pattern of melting curve was suggestive of the existence of a potential mutation or single nucleotide polymorphism (SNP), the sample was diluted 1/20 on sterile water and sequenced using BigDye Cycle sequencing kit (Applied Biosystems). Any variant confirmed by sequencing was then genotyped by TaqMan technology in a 7900HT Fast Real-Time PCR system, under conditions recommended by the manufacturer (Applied Biosystems, Foster City, CA, USA).

2.5. Statistical analysis

Statistical analyses were performed using SPSS 15.0. Linkage disequilibrium values (D') and Hardy–Weinberg equilibrium were tested with Haploview 4.0 software. Chi-Square test was used to compare allele and genotype frequencies. Mann–Whitney U test was used to compare MSRV-like env relative expression between two groups. Statistically
3. Results

In order to explore the genomic insertions of HERV-W env that match with the published MSRV env sequence (AF331500), homologies were searched in the human genome assembly GRCh37.p5 using BLAST (see Methods). The overall 261 sequences resulting from this query were searched in the human genome assembly GRCh37.p5 using BLAST match with the published MSRV sequence (AF331500), homologies were searched in the human genome assembly GRCh37.p5 using BLAST (see Methods). The overall 261 sequences meeting these criteria and were selected for further analyses (Supplementary Fig. 1). From these genomic insertions of MSRV-like env, the two of them with ORFs around 10% shorter than the physiological syncytin (1617 bp) were: one located in chromosome X (1425 bp) and the other in chromosome 20 with 1416 bp (Fig. 1). The region of chromosome X which would encode a truncated syncytin protein has been previously studied [26] and in the present work we focused on the other plausible candidate in chromosome 20 (for protein alignments see Supplementary Fig. 2).

A preparative PCR amplification step from genomic DNA was conducted to isolate the specific MSRV-like env copy in chromosome 20 (Fig. 2), yielding a 4232 bp product used as template to perform HRM analysis in overlapping amplicons (for primers used and PCR conditions see Supplementary Table 1). Whenever a potential variant was detected in an initial screening of 100 samples, it was validated by sequencing and then analyzed by TaqMan technology in the overall 668 MS patients and 678 healthy controls. Two SNPs were identified following this procedure: T/C and T/G changes at positions 221 and 285 from the beginning of the amplification product (D' = 35%, NCBI_ss# SNP1: 974293065 and SNP2: 974293066) (Fig. 3).

Genotyping success was over 95% for both patient and control cohorts and no departure from Hardy–Weinberg equilibrium was observed. However, no significant differences were found in the allelic or genotypic frequencies between MS patients and ethnically matched controls ($\chi^2$ test: p > 0.05, Table 1). Moreover, no significant associations (Mann–Whitney U test: p > 0.05) were found between alleles or genotypes of these SNPs and the RNA expression of MSRV env specifically derived from the chromosome 20 copy (Fig. 4).

4. Discussion

In the present work, we aimed to shed some light on the genomic origin of the reported MSRV env protein previously identified as MS pathogenic co-factor. Specifically, we pursued to elucidate the putative role in MS risk of the insertion of MSRV-like env in chromosome 20 as one of the plausible candidates by size, together with the insertion in chromosome X already studied by our group [26]. Both chromosomal origins showed similar lengths of their respective ORFs, 10% shorter than the one measured for syncytin [29] and could putatively originate functional proteins. The ENV protein has been recently evidenced with post-mortem MS brain lesions [15,30–32]. MSRV ENV induces human monocytes to produce proinflammatory cytokines through engagement of CD14 and TLR4 [19]. Kremer and collaborators demonstrated that the ENV protein is present in close proximity to TLR4-expressing oligodendroglial precursor cells adjacent to MS lesions [33]. In addition, MSRV ENV triggers a maturation process in human dendritic cells driving a Th1-like differentiation [19].

We recently reported that the MSRV-like env copy number is increased in PBMCs of MS patients, as an elevated MSRV-like env load was observed in MS patients compared to controls ($p = 4.15 \times 10^{-7}$) [34]. Moreover, in our cohort the transcription levels of MSRV-like env were higher in MS patients than in controls (Mann–Whitney U test: $p = 0.004$) [26], consistent with previous findings [24,35]. Nonetheless, knowledge is still lacking about which chromosomal location(s) harboring specific retroviral copy(ies) originate(s) these induced levels of transcription. Once the RNA expression in PBMCs was significant differences were considered when the p-value was lower than 0.05.

### Table 1

Allelic and genotypic frequencies of both SNPs identified in the HERV-W copy of chromosome 20 in MS patients and controls.

<table>
<thead>
<tr>
<th>SNP1</th>
<th>MS</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT</td>
<td>495</td>
<td>75</td>
</tr>
<tr>
<td>TC</td>
<td>153</td>
<td>23</td>
</tr>
<tr>
<td>CC</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>T</td>
<td>1143</td>
<td>86</td>
</tr>
<tr>
<td>C</td>
<td>181</td>
<td>14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SNP2</th>
<th>MS</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT</td>
<td>652</td>
<td>99</td>
</tr>
<tr>
<td>TG</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>GG</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>G</td>
<td>9</td>
<td>1</td>
</tr>
</tbody>
</table>

*Fig. 4.* Relative expression of HERV-W env copy on chromosome 20 in 52 MS patients and 53 healthy blood donors (A) and stratified by genotype of the identified SNPs (B and C). Lines inside the boxes represent medians, whiskers represent standard deviations.
validated (data not shown), we analyzed the insertions encoding modified envelope proteins: those derived from chromosomes X and 20. It has been demonstrated that MSRV-like env expression in brain phagocytes within MS lesions is consistent with the observed expression in MS PBMCs, supporting the rationale for its ex vivo detection in peripheral blood [30]. Nonetheless, our findings regarding the transcriptionally active long ORF derived from chromosome 20 dismiss this copy as a predominant origin of the MSRV pathogenic effect.

Previous results of GWAS in MS demonstrated associations with three loci within chromosome 20 including the following genes: CD40, TNFRSF68 and CYP24A1 [2], the latter at 1 Mb of the HERV-W insert studied in the present work. Considering that repetitive sequences of the genome have not been explored in the GWAS, this seemed an adequate analysis to advance in the unresolved issue of the missing heritability of this complex disease [4]. Moreover, one of the proposed co-factors repeatedly involved in the pathogenesis of MS is the Epstein–Barr virus, which was reported to activate in vitro the HERV-W family of human endogenous retroviruses [36]. The specific PCR amplification of the genomic region of chromosome 20 including the MSRV-like env sequence and subsequent HRM analysis led to the identification of two polymorphisms within this repetitive insertion. However, the performed case-control study did not detect significant differences in allelic or genotypic frequencies for these SNPs. Therefore, the results obtained in this study do not encourage further investigation of this MSRV-like env insertion as responsible of part of the genetic risk in MS patients. Despite the mentioned similar size of the two MSRV-like env copies from chromosomes 20 and X, apparent differences have been found between them. One polymorphism identified in chr. X, rs6622139*T, was associated with higher MSRV-like env expression (Mann–Whitney U test; p = 0.003), while the two polymorphisms found in chromosome 20 did not show evidence of association with chromosome 20-specific RNA expression levels. Additionally, rs6622139 was associated in women with MS susceptibility and severity [26]. These results point to the insertion in chromosome X and not the one in chromosome 20 as involved in the origin of MSRV ENV.

Suggestive of its potential pathogenic activity, significant decreases in seroreactivities to HERV-W envelope antigens, closely linked to efficacy of interferon-beta therapy were previously demonstrated [37] and recently, it has been reported that after one year therapy, natalizumab inhibits the plasma-membrane levels of the ENV protein in MS patients [38]. The challenge is now to turn the increasing knowledge in HERV-W to the clinical setting, as it has proven helpful in predicting therapy outcomes [34]. Supplemental data to this article can be found online at http://dx.doi.org/10.1016/j.bbbcl.2015.02.002.

Transparency document

The Transparency document associated with this article can be found, in the online version.

Acknowledgments

This work was supported by: Instituto de Salud Carlos III–Fondo Europeo de Desarrollo Regional (Feder), Fondo de Investigaciones Sanitarias FIS (PI13/0879) and Fundaciones: Genzyme (to R.A.-L.); Lai (to R. A.) and Mutua Madrileña (to R. A.-L). R.A.-L. is recipient of a research contract of the Instituto de Salud Carlos III–Feder (CP07/00273). M.A.G.-M. is recipient of a technician contract from the “REEM: Red Española de Esclerosis Múltiple” (RETICS-REEM RD12/0032/009; www.reem.es). B.H. is recipient of a PhD scholarship from the “Fondo de Investigaciones Sanitarias” (F11/00560). J.V. is recipient of a contract from the “Ministerio de Economía y Competitividad” (PTA2011-6137-1) and E.U. works for the FIB-Hospital Clínico San Carlos-IdISSC. The authors do not disclose any conflict of interest.

References

[34] M. Garcia-Montojo, et al., The DNA copy number of human endogenous retrovirus-W (MSRV-type) is increased in multiple sclerosis patients and is influenced by gender and disease severity, PLoS One 8 (2013) e56323.

